

2020 Census Research and Testing 2014 Census Test - Analysis of Administrative Record Usage

A New Design for the 21st Century

Issued March 25, 2016

Version 1.5

Prepared by Decennial Census Management Division



[This page intentionally left blank]

Contents

List of Tables	ii
Executive Summary	1
1. Introduction	1
1.1 Scope	1
1.2 Intended Audience	1
2. Background	2
3. Methodology	3
3.1 Research Questions	5
3.2 Assumptions	5
3.3 Office of Management and Budget Clearance	6
3.4 Schedule	6
4. Limitations	7
5. Results	8
5.1 The Control Panel and the Reduced Contacts without AR Panel Analysis	9
5.2 Reduced Contacts with Full AR Removal Panel Analysis	23
5.3 Adaptive Design with Hybrid AR Removal Panel Analysis	24
5.4 ‘What If?’ Scenarios	28
6. Related Evaluations, Experiments, and/or Assessments	39
7. Dependencies	39
8. Conclusions and Recommendations	39
8.1 Conclusions	39
8.2 Program-Level Recommendations	40
8.3 Project-Level Recommendations	40
9. Knowledge Management Resolutions	41
10. Acknowledgements	41
11. References	41
12. Document Logs	41
12.1 Sensitivity Assessment	41
12.2 Review/Approval	42
12.3 Version History	42

List of Tables

Table 1: AR Initial Flags by Panel (NRFU Universe)..... 4

Table 2: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel..... 9

Table 3: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel..... 9

Table 4: Resolved Housing Unit Type Status for Cases Assigned a False Positive AR Vacant Flag by Panel..... 10

Table 5: Housing Unit Type for Resolved Cases Assigned a False Positive AR Vacant Flag by Panel 10

Table 6: Did Household Say No One Lived There on Census Day?..... 11

Table 7: Resolved Household Size Status on Census Day for AR Vacant and NRFU Occupied Cases by Panel..... 11

Table 8: Household Size on Census Day for Resolved AR Vacant and NRFU Occupied Cases by Panel..... 12

Table 9: UAA Reason Codes for False Positive and True Positive AR Vacant Cases (Control Panel and the Reduced Contacts without AR Panel Combined) 12

Table 10: Vacant Flag Assignment by Panel 13

Table 11: Resolved Housing Unit Status for Cases Not Assigned a Vacant Flag by Panel 13

Table 12: Housing Unit Status for Resolved Cases Not Assigned a Vacant Flag by Panel 13

Table 13: Reasons for Missed Vacant Flags for Vacant Interviews by Panel 14

Table 14: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel..... 14

Table 15: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel..... 15

Table 16: Resolved Housing Unit Status by Household Composition for AR Occupied Cases by Panel..... 15

Table 17: Housing Unit Status by Household Composition for Resolved AR Occupied Cases by Panel..... 16

Table 18: Resolved Population Count Status for True Positive AR Occupied Cases by Panel..... 16

Table 19: Match Status of Population Counts for Resolved True Positive AR Occupied Cases by Panel. 17

Table 20: Resolved Population Count Status by Household Composition for True Positive AR Occupied Cases by Panel 17

Table 21: Match Status of Population Counts by Household Composition for Resolved True Positive AR Occupied Cases by Panel 18

Table 22: Population Count Comparison for Resolved, Non-Matched True Positive AR Occupied Cases by Panel..... 18

Table 23: Population Count Comparison by Household Composition for Resolved, Non-Matched True Positive AR Occupied Cases by Panel..... 19

Table 24: Population Count Comparison for Resolved True Positive AR Occupied Cases by Panel 19

Table 25: Comparison of AR and NRFU Population Counts for Resolved Cases by Panel	20
Table 26: Resolved Housing Unit Type Status for Cases Assigned a False Positive AR Occupied Flag by Panel.....	20
Table 27: Housing Unit Type for Resolved Cases Assigned a False Positive AR Occupied Flag by Panel.....	21
Table 28: Did Household Say Anyone Lived there on Census Day?	21
Table 29: Occupied Flag Assignment By Panel	21
Table 30: Resolved Housing Unit Status for Cases Not Assigned an Occupied Flag by Panel	22
Table 31: Housing Unit Status for Resolved Cases Not Assigned an Occupied Flag by Panel	22
Table 32: Reasons for Missed Occupied Flags for Occupied Interviews by Panel	22
Table 33: Resolved Housing Unit Status of Cases Not Removed in the Reduced Contacts with Full AR Removal Panel	23
Table 34: Housing Unit Status of Resolved Cases Not Removed in the Reduced Contacts with Full AR Removal Panel	23
Table 35: Reasons for Missed Vacant Flags for Vacant Interviews for the Reduced Contacts with Full AR Removal Panel	24
Table 36: Reasons for Missed Occupied Flags for Occupied Interviews for the Reduced Contacts with Full AR Removal Panel	24
Table 37: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag in the Adaptive Design with Hybrid AR Removal Panel	25
Table 38: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag in the Adaptive Design with Hybrid AR Removal Panel	25
Table 39: Resolved Housing Unit Status by Household Composition for AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel	25
Table 40: Housing Unit Status by Household Composition for Resolved AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel	26
Table 41: Resolved Population Count Status for True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel	26
Table 42: Match Status of Population Counts for Resolved True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel	26
Table 43: Resolved Population Count Status by Household Composition for True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel	27
Table 44: Match Status of Population Counts by Household Composition for Resolved True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel.....	27
Table 45: Population Count Comparison for Resolved, Non-Matched True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel.....	27

Table 46: Population Count Comparison by Household Composition for Resolved, Non-Matched True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel.....	28
Table 47: Population Count Comparison for Resolved True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel.....	28
Table 48: AR Initial Flags by Panel: Alternate AR Removal Scenario Allowing Only a UAA Vacant Reason Code for Flagging AR Vacant Removals.....	29
Table 49: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel.....	29
Table 50: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel.....	30
Table 51: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel.....	30
Table 52: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel.....	30
Table 53: Percent Change in AR Removals by Panel.....	31
Table 54: Percent Change in NRFU Workload by Panel.....	31
Table 55: AR Initial Flags By Panel: Alternate AR Removal Scenario Allowing Only Specific Household Compositions for Flagging AR Occupied Removals.....	32
Table 56: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel.....	32
Table 57: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel.....	32
Table 58: Percent Change in AR Removals by Panel.....	33
Table 59: Percent Change in NRFU Workload by Panel.....	33
Table 60: AR Flags By Panel: Alternate AR Removal Scenario Requiring a Match Between IRS 2012 and the 2013 Households for Flagging AR Occupied Removals	34
Table 61: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel.....	34
Table 62: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel.....	35
Table 63: Resolved Housing Unit Status of Cases Assigned an AR Occupied Flag by Panel	35
Table 64: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel.....	35
Table 65: Percent Change in AR Removals by Panel.....	36
Table 66: Percent Change in NRFU Workload by Panel.....	36
Table 67: AR Flags by Panel: Alternate AR Removal Scenario Using the Selective Service, HUD TRACS and Indian Health Service Files	37
Table 68: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel.....	37
Table 69: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel.....	37
Table 70: Resolved Housing Unit Status of Cases Assigned an AR Occupied Flag by Panel	38
Table 71: Housing Unit Status of Resolved Cases Assigned an AR Occupied Flag by Panel	38
Table 72: Percent Change in AR Removals by Panel.....	38
Table 73: Percent Change in NRFU Workload by Panel.....	39

Executive Summary

The 2014 Census Test was the first 2020 Census Test in which administrative record information was utilized to reduce the number of contacts necessary to follow up at units that did not respond to the initial mailing. The administrative record applications in this test were able to build on the pilot test done in the 2013 Census Test where 2010 Nonresponse Followup (NRFU) cases were tested. This report summarizes how cases were identified during the production part of the test and analyzes the findings.

The production processing utilized administrative record information from the following sources: Internal Revenue Service Individual Income Returns, Center for Medicare and Medicaid Services Medicare Enrollment Database, Social Security Administration Numerical Identification file and United States Postal Service Undeliverable-As-Addressed information. The production processing identified occupied and vacant units from two of the four panels in the test.

Here are some of the key findings related to the identification of vacant units:

- The administrative record processing was able to identify 4.8 percent of the NRFU cases in the Reduced Contacts with Full Administrative Record Removal Panel and 4.1 percent of the NRFU cases in the Adaptive Design with Hybrid Administrative Record Removal Panel as being vacant based on their administrative record information.
- This analysis utilized the Control Panel and the Reduced Contacts without Administrative Record Panel to compare identified administrative record vacant cases to the results based on the census interview. For cases that were determined to be administrative record vacant, the results show agreement on a vacant status 53.3 percent and 54.2 percent of the time in the Control Panel and the Reduced Contacts without Administrative Record Panel, respectively. The results show agreement on unoccupied status 69.9 percent and 75.4 percent of the time in the Control Panel and the Reduced Contacts without Administrative Record Panel, respectively.
- Using administrative record information, the processing assigned a vacant status to a NRFU occupied unit 30.1 percent and 24.6 percent of the time in the Control Panel and the Reduced Contacts without Administrative Record Panel, respectively. For these occupancy status disagreements in the control panel, about half them were instances where the occupied NRFU results had conflicting answers in the interviews, or the processing was unable to determine a population count.

Here are some of the key findings related to the identification of occupied units:

- The administrative record processing was able to identify 57.5 percent of the NRFU cases in the Reduced Contacts with Full Administrative Record Removal Panel, and 56.1

percent of the NRFU cases in the Adaptive Design with Hybrid Administrative Record Removal Panel as being occupied based on their administrative record information.

- Similarly, for occupied units, this analysis utilized the Control Panel and the Reduced Contacts without Administrative Record Panel to compare administrative record occupied cases to the results based on the census interview. For administrative record occupied cases, the results show agreement on occupancy status over 90 percent of the time across the two panels. While this was high, the population count agreement was only 54 percent in the Control Panel. Further analysis showed that the population count agreement was higher for single adult with no children, two adult with children and two adult without children household compositions. This finding suggests giving priority to these household compositions when determining when to use administrative records.

The last part of our analysis looked at several scenarios of using administrative records differently than used in production. These scenarios were examined using the two panels that did not remove cases from the workload based on administrative record information. This analysis used the production data sources in different ways plus including additional sources as well. Additional sources examined were the Selective Service System Registration System, Department of Housing and Urban Development Tenant Rental Assistance Certification System and Indian Health Service Patient Registration System information. Here are some of the key findings.

- The first “what-if” scenario examined if the processing was stricter about the Undeliverable-As-Addressed Nixie reasons codes used. Specifically, we examined what would occur if only the vacant reason code from the first mailing was used. The result was 58.4 percent true positive rate when identifying administrative record vacant units. In production, this value was only 53.8 percent. However, by using this stricter criterion, the overall NRFU workload was increased by 1.8 percent.
- The second “what-if” scenario examined what would occur if the processing was stricter about the types of households used to determine administrative record occupied cases. Comparisons showed that the household compositions of single adult with no children present, two adults without children present, and two adults with children present had higher population count agreements. This analysis restricted the production data to only these three compositions to make administrative record households. The results showed a 93.2 percent true positive rate when identifying administrative record occupied cases. In production, this value was 94.1 percent. The results showed a 64.6 percent population count match rate between our administrative record occupied cases and the census results in this scenario, compared to just a 56.1 percent agreement in production. However since this approach was stricter, it resulted in fewer administrative occupied records being identified, and thus would increase for the NRFU fieldwork. Assuming a full administrative record removal strategy, this would have resulted in a 36.4 percent increase in the NRFU workload.
- A third “what if” scenario expanded our production rules to include information from three additional federal sources, the Selective Service System Registration System,

Housing and Urban Development Tenant Rental Assistance Certification System and the Indian Health Service Patient Registration System. This scenario resulted in a 93.9 percent true positive rate when identifying administrative record occupied cases. In production, this value was 94.1 percent. Adding in these three sources would have reduced the NRFU workload, assuming a full administrative record removal strategy by 2.5 percent. The count comparison was similar to the production removal.

Based on this analysis, we have the following recommendations:

- We recommend continuing to research approaches to identify vacant and occupied units using administrative record and third-party information.
- We also recommend testing using predictive modeling approaches in addition to rule-based approaches to identify vacant and occupied units using administrative record and third-party information.

1. Introduction

To meet the strategic goals and objectives of the 2020 Census, the Census Bureau must make fundamental changes to the design, implementation, and management of the decennial census. These changes must build upon the successes and address the challenges of the previous censuses while also balancing challenges of cost containment, quality, flexibility, innovation, and disciplined and transparent acquisition decisions and processes.

For the nonresponse operation, the 2014 Census Test compared cost and data quality across several strategies aimed at reducing the cost of implementing the operation. This included:

- Modifying contact strategies by reducing the number of contacts, making more use of the telephone for contact, and applying adaptive design methods to manage the work in the field.
- Using administrative records to remove cases by assigning a vacant status or to enumerate the housing units.

1.1 Scope

The goal of this study was to evaluate the use of administrative records (ARs) to remove cases from the Nonresponse Followup (NRFU) workload. AR information was employed to assign a housing unit status and a household population to units that would have otherwise been sent to NRFU, thereby reducing the total NRFU workload. We flagged occupied housing units for potential AR enumeration using two AR sources: IRS Individual Taxpayer Returns (IRS 1040) and Center for Medicare and Medicaid Services Medicare Enrollment Database (CMS MEDB). The Social Security Number Identification File was utilized to assign protected identification keys (PIKs) to each record. Rastogi and O'Hara (2012) describe these two data sources in the Census Match Study. The IRS 1040 and MEDB are data sources from other federal agencies. For the 2014 Census Test, we also used Undeliverable-as-Addressed (UAA) Nixie information from the United States Postal Service (USPS) obtained from the first mailing attempt.¹ We flagged vacant housing units for potential AR removal if the prenotice first mailing attempt resulted in an UAA designation indicating vacancy, or any of the other reason codes listed in Section 3.

1.2 Intended Audience

This report is intended for use by team members and Census Bureau management to develop plans for future testing and decision making for the 2020 Census. It is assumed that the reader has a basic understanding of the Census.

¹ The USPS classified mail that cannot be delivered by postal mail carriers as UAA, and such mail are sent into a special operation.

2. Background

In the 2014 Census Test, the NRFU operation had four panels that employed different contact strategies and different ways of using AR files, including a control panel that used no ARs. One panel had an adaptive contact strategy while the others had a fixed contact strategy.

- The Control Panel mimicked the 2010 Census NRFU contact strategy as closely as possible. A maximum of six contact attempts was permitted with a proxy response permitted only after the maximum attempts to interview a household member had failed. Addresses where interviewers could not obtain a proxy, or the proxy responses were not data-defined, received count imputations. The contact strategy was fixed for all households and did not use ARs in any way. This panel served as a control for comparing the results of the other panels.
- The Reduced Contact Strategy, no Administrative Records Use Panel (Reduced Contacts without AR Panel) had a strategy that permitted fewer contact attempts than allowed in the Control Panel. The reduced contact strategy called for the first contact attempt to be a personal visit. If the interviewer could not contact a household member or resolve the address as vacant or non-existent on the first attempt, then the second attempt was made by telephone. The telephone number was obtained by matching the address to a commercial database. If the attempt to contact by telephone was unsuccessful, the interviewer made a third attempt, again by personal visit. If the interviewer was still unable to contact a household member on this third attempt, then the interviewer could utilize proxy respondents. Addresses where interviewers could not obtain a proxy, or the proxy responses were not data-defined, received count imputations. This approach used a fixed contact strategy for all households and did not use ARs in any way. This panel serves to evaluate the reduced contact strategy without any confounding information from the use of ARs.
- The Reduced Contact Strategy with Administrative Records Panel (Reduced Contacts with Full AR Removal Panel) employed the use of ARs. First, addresses identified by ARs as vacant or having ARs of sufficient quality to use for enumeration were removed from the NRFU field workload. The fixed reduced contact strategy described for the Reduced Contacts without AR Panel was implemented for all remaining addresses. Count imputations were created for addresses where interviewers could not obtain a proxy or the proxy responses were not data-defined. This panel served to evaluate extensive use of ARs.
- The Adaptive Design with Administrative Records Panel (Adaptive Design with Hybrid AR Removal Panel), employed a hybrid AR removal strategy for using ARs to remove cases from the NRFU workload as well as an adaptive design strategy to reduce the number of contacts for each case. Initially, housing units identified as vacant using ARs and postal service Undeliverable-As-Addressed (UAA) information were removed from the workload. For housing units that had not been removed as vacant or completed by Computer Assisted Telephone Interviewing, and for which we had ARs indicating an occupied status, enumerators made only one personal visit attempt. No proxies were

allowed for these cases. Cases unresolved after one personal visit were enumerated using ARs. The use of ARs to remove vacant units prior to interviewing and to enumerate occupied units after one personal visit attempt is the hybrid AR removal strategy, and is the focus of the analysis in this document. See Poehler et al (2014) for more information about the adaptive design component of this panel.

3. Methodology

Creation of Person Level AR File

For the IRS Individual Taxpayer Returns, we obtained the 2013 IRS 1040 Individual Tax Returns for the first 26 weeks of the year. For the Medicare Enrollment Database, we obtained the 2012 and 2013 files. We converted each file into a person-level file where a unique combination of Master Address File Identification number (MAFID) and PIK define a person.

A PIK may be duplicated across multiple MAFIDs. We retained only records where the MAFID and PIK were non-blank. Units containing unvalidated person records, those where a PIK could not be assigned because of insufficient information, were removed from the file. For the MEDB files, we also removed from consideration any records that had a non-blank beneficiary date of death.

Assignment of Occupied Housing Unit Flag

We applied two rules to flag housing units as occupied. The two rules were:

- There must not be any unvalidated person records for the housing unit of the given AR source. Each person in the housing unit needed to be assigned a PIK.
- For the MEDB source, the housing unit must have the same set of people based on PIKs on the 2012 and the 2013 version of the source. Each PIK in the unit on the 2012 file must be in the unit on the 2013 file and vice versa.

Any unit that passed these two rules for a given AR source was flagged as occupied due to that source. All persons in the given AR source were used to generate a household population count and define certain characteristics of the housing unit. Some units were flagged as occupied due to multiple sources. These sources identified different sets of persons. All of the persons from all sources were retained. If any persons overlapped between multiple sources, they were only counted once in the housing unit.

Assignment of Vacant Housing Unit Flag

For the test, we used UAA information from the USPS following the first mailing. Units flagged as vacant were UAA with the following reason codes:

- Attempted - Not Known - Delivery attempted, addressee not known at place of address (ANK)
- Deceased (DEC)
- In Dispute (DIS)
- Illegible (ILL)

- Refused - Addressee refused to accept mail or pay postage charges on it (REF)
- Unclaimed - Addressee abandoned or failed to call for mail (UNC)
- Not Deliverable as Addressed - Unable to Forward - Mail undeliverable at address given; no change-of-address order on file; forwarding order expired; forwarding postage not guaranteed by sender or addressee; or, mail endorsed with sender's instructions DO NOT FORWARD (UTF)
- Vacant - House, apartment, office, or building not occupied (Use only if mail addressed "Occupant.") (VAC)

Formation of the Analysis Universe

As described above, each sample unit was eligible to receive occupied and vacant flags, regardless of their panel. A flag was set for an occupied or vacant unit as defined by the conditions above. If a unit was assigned both an occupied flag and a vacant flag, this unit fell into the 'No AR Identification' category. Because of the inconsistent information between AR sources and UAA information, we sent these cases to NRFU to be resolved. Table 1 shows the initial AR flag assignments of the NRFU universe across all panels.

Table 1: AR Initial Flags by Panel (NRFU Universe)

Panel	# Units N	Initial Flags					
		No AR Identification		AR Occupied		AR Vacant	
		%	SE	%	SE	%	SE
Control	13,253	43.5	3.3	53.0	3.1	3.5	0.6
Reduced Contacts without AR	12,553	38.9	1.5	55.8	1.6	5.3	0.7
Reduced Contacts with Full AR Removal	8,101	37.7	3.0	57.5	3.1	4.8	0.9
Adaptive Design with Hybrid AR Removal	12,340	39.8	2.1	56.1	2.3	4.1	0.6
Total	46,247	40.2	1.3	55.4	1.3	4.4	0.3

Assignment of Household Composition

Many of the tables in Section 5 are broken down by household composition, which we defined by the number of adults and children in each household. The formation of a household roster from the IRS 1040 and MEDB sources is described in Section 3. In addition, we used age to distinguish persons as either adults (18 or more) or children (under 18). Age was taken from the Social Security Administration Number Identification File. Persons with missing values of age were coded as children.

Assignment of Housing Unit Status and Household Population Count

The NRFU Design and Operations team provided both the housing unit status, which designates units as occupied, vacant, or non-existent, as well as the household population count, to the Administrative Records Modeling team. Since this was a test, the housing unit status and population count processing did not include all of the processing steps involved in the 2010 Census processing.

Calculation of Standard Errors

For each table presented in this document, estimates are accompanied by their standard errors. These standard errors were calculated using a Taylor Series Linearization method. Since our sample design included clustering, and our primary sampling units, in this case block groups, were randomly assigned to the four panels, our standard errors took into account the variance among block groups.

3.1 Research Questions

- How many vacant units could be identified in production based on AR information?
- How many occupied units could be identified in production based on AR information?
- How does the AR information compare to Nonresponse Followup results?
- How do the AR usage results change for different assumptions of AR usage?

3.2 Assumptions

Research and planning for the 2020 Census has been focused on major innovations to the design of the census, oriented around the major cost drivers of the 2010 Census. Identification of those cost drivers led to four major design principles/research tracks: Reengineering Address Canvassing, Optimizing Self-Response, Utilizing Administrative Records, and Reengineering Field Operations. The overarching assumptions within each design principle, see Bishop (2014), must be proven or revised in preparation for the 2020 Census design decisions:

Reengineering Address Canvassing:

- In-field address canvass only 20 percent of total housing units
- Eliminate early-opening local census offices (manage from Regional Census Centers)
- Redesign the training strategy to reduce enumerator training hours by 35 percent
- Reduce the number of crew leader assistants by 50 percent
- Establish a training pay rate of \$1.50 lower than the production pay rate

Optimizing Self-Response:

- Promote for internet self-response from 55 percent of the population
- Mail paper questionnaires to only a targeted 20 percent of nonrespondents
- Achieve a reduction in paper data capture operations and infrastructure as compared to the 2010 Census

Utilizing Administrative Records:

- Reduce the total NRFU workload by 11 percent through the removal of vacants and deletes
- Reduce the total number of local census offices by 12 percent through the removal of vacants and deletes

- Eliminate Coverage Followup and Vacant/Delete Operations
- Reduce the total number of NRFU visits

Reengineered Field Operations:

- Increase NRFU productivity by 20 percent with automation
- Remove late responses from the NRFU workload
- Reduce the total number of local census offices by 5 percent
- Reduce the total square footage of local census offices by 70 percent
- Eliminate Crew Leader assistants
- Reduce the number of clerical staff by 20 percent with automation
- Redesign the training strategy to reduce enumerator training hours by 35 percent
- Establish a training pay rate \$1.50 lower than the production pay rate
- Allow seventy-five percent of enumerators to bring their own device (BYOD)
- Reduce the phone/personal visit contact cycle relative to the 2010 Census
- Use routing and dynamic case management to allocate resources efficiently

From the above research areas, the removal of vacants and deletes by utilizing administrative records is relevant for this evaluation study.

3.3 Office of Management and Budget Clearance

This research project, the 2014 Census Test, was covered under the Office of Management and Budget (OMB) clearance number 0607-0979 that expired on 5/31/2015.

3.4 Schedule

Below was the completion schedule for this analysis.

Analysis Report Milestone Schedule

Activity	Start	End
Develop preliminary results for internal customers	8/25/2014	11/7/2014
Prepare initial Draft Project Report	10/9/2014	11/7/2014
Distribute initial draft to team for review	11/10/2014	11/14/2014
Incorporate comments	11/17/2014	11/21/2014
Brief final draft product to team Program Manager	11/24/2014	11/26/2014
Incorporate comments	11/28/2014	12/2/2014
Brief final draft product to the 2020 R&T Strategies Group		12/3/2014
Prepare final product	12/4/2014	12/9/2014
Brief final product to the Decennial Leadership Group		12/10/2014
Prepare final product for base-lining, incorporating and addressing all comments	12/11/2014	
Send final product to 2020 Program Management mailbox for base-lining		

4. Limitations

Because this test was conducted in a specific geographic site, the results are only representative of the households and persons in the specific site. The results cannot be generalized to any populations outside the specific site (e.g., national population, populations in other metropolitan areas, populations in other types of geographic areas, etc.). Therefore, the results of this test can only be used to draw general conclusions regarding the treatments being tested, such as relative differences between treatments, and cannot be used to estimate or predict specific results we might see in the 2020 Census.

Comparisons between the control panel and 2010 Census results may be limited based on differences in procedures used in the Control panel and the 2010 Census. In the Control panel, telephone numbers were provided to enumerators to help with data collection based on a telephone number lookup with ARs. In 2010, no lookup was performed. Additionally, interviews were conducted using an automated, handheld instrument in this test, while enumerators used paper questionnaires in the 2010 Census. Lastly, in this test, the “Notice of Visit” was left at housing units when someone was not home (and other situations) directed respondents to go to the Internet or to call the Telephone Questionnaire Assistance center to complete the census questionnaire. In the 2010 Census, the “Notice of Visit” directed respondents to call the enumerator or the Local Census Office to complete the interview or schedule an appointment.

One limitation was that the census test had an issue with the four mailings to the address. The team was looking to obtain the UAA information about each of those mailings. The UAA information is obtained based on a barcode that is associated with each mail piece. Each barcode is associated with a Census Identification number. During the mailing, it was identified that barcodes from earlier mailings were reused in later mailings. Because of this, UAA results could not be obtained for later mailings. Since barcodes were associated with multiple mailings, it was not possible after a certain point in time to associate the UAA results with the correct Census ID since it was linked to multiple mail pieces.

This site test had a Census Day of July 1, 2014. Since the majority of IRS data filings were from February through April, there was a time lag between the administrative data residence and their Census Day residence. This lag will be much smaller in 2020, as the 2020 Census Day is April 1, 2020. Medicare data also had a time lag, but required a two-year match of PIKs to qualify as an AR occupied identified case. This increased the likelihood that we had non-moving household members, and thus, that the AR and census data would provide the same address information despite the lag. Analysis here may show larger differences between administrative record and census results than may be observed if Census Day was April 1st for this test.

The NRFU panel design was not a full factorial design, thus comparisons to be able to separate out effects due the various AR identification strategies or contact strategies is limited. As a result, we focused on the administrative record comparisons

Persons with missing values of age were coded as children in the formation of the household composition variable. Other approaches may more accurately determine whether these persons were adults or children. This will be explored in future research. Future work will also investigate the implications of using different compositions. Comparisons to census household counts were based on one set of rules to determine housing unit status and household counts based on the response and auxiliary data available. Different rules could have resulted in different comparison results than those shown here.

All tables presented in this document have percentages and standard errors rounded to the nearest tenth. Therefore, if the percentages do not sum to 100.0 percent, this is due to rounding. In addition, all tables, unless otherwise noted, are representing the NRFU universe or a subset thereof.

5. Results

The goal of this study was to evaluate the use of ARs to remove cases from the NRFU workload. We wanted to learn how ARs could best be implemented in future tests as well as the 2020 Census. We performed two types of analysis with the 2014 results. First, we completed an analysis evaluating the methodology that was used in the 2014 Census Test. To do this, we assessed the agreement with census fieldwork for ARs identified as occupied and vacant cases. The second type of analysis offered ‘what if’ scenarios. This section looks at how the results would have been different if we would have altered the methodology for identifying occupied and vacant cases.

In this both types of analysis, we were comparing the status assigned by ARs against the results seen in the field. The Control Panel and the Reduced Contacts without AR Panel completed interviews without the use of ARs. As a result, we can compare the NRFU interview results with what would have been obtained had ARs been used.

The Reduced Contacts with Full AR Removal Panel used ARs to allocate vacant and occupied cases before any NRFU contacts were made. The analysis for this panel therefore revolved around occupied and vacant cases we found after going into the field. It allowed us to examine why these were not identified. Finally, the Adaptive Design with Hybrid AR Removal Panel used ARs to allocate vacant cases before any NRFU contacts were made. ARs were also used to allocate occupied cases in this panel after one NRFU contact was made.

5.1 The Control Panel and the Reduced Contacts without AR Panel Analysis

This analysis was divided into two parts. The first part analyzed addresses removed as vacant. The second part analyzed addresses removed as occupied. Since ARs were not used to remove workload in either the Control Panel or Reduced Contacts Without AR Panel, this section shows results for both panels.

5.1.1 Comparing AR Vacant cases to NRFU Enumerations

We were interested in how often our AR assignment of a vacant status matched to the NRFU field status, as well as how often it did not match. To understand the validity of our assignment method, Tables 2 and 3 show the resolved housing unit status and the final housing unit status, respectively, for cases in the Control Panel and the Reduced Contacts without AR Panel.

Table 2 shows the assignments of vacant status agreed 53.3 percent and 54.2 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. Since we classified these units as vacant, and they were determined to be so in the field, these units can be thought of as true positives. Combining the vacant and non-existent results, assignment of unoccupied status agreed 69.9 percent and 75.4 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively.

Table 3 shows that AR vacant status was determined to be occupied in NRFU 30.1 percent and 24.6 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. Additionally, cases assigned an AR vacant status were determined to be a NRFU delete 16.7 percent and 21.2 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. Both of these differences can be thought of as false positives, i.e., the cases for which we assigned a vacant status, but were occupied or delete in NRFU.

Table 2: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	462	79.2	2.4	20.8	2.4
Reduced Contacts without AR	665	71.6	4.2	28.4	4.2

Table 3: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	366	30.1	4.3	53.3	4.3	16.7	5.1
Reduced Contacts without AR	476	24.6	3.6	54.2	7.8	21.2	8.4

5.1.1.1 False Positive Analysis

This part analyses the false positives. This section examined further why cases that were assigned a vacant status based on administrative record information was determined to be

occupied or delete in the field. This section focused on the occupied NRFU cases since those differences have population implications. Were these false positives occurring for specific types of units? Were these units assigned a vacant code at any point in the enumeration process, but then ultimately resolved as occupied? The following tables attempt to answer these questions.

Housing Unit Type

Table 4 shows how often the housing unit type was resolved for these false positive cases assigned an occupied status in NRFU. In the Control Panel, 89.1 percent of the units had a resolved housing unit type. In the Reduced Contacts without AR Panel, this value was 84.6 percent. The result was lower than expected for the control panel. Since the Reduced Contacts without AR panel had fewer contacts, some amount of unresolved was expected.

Table 4: Resolved Housing Unit Type Status for Cases Assigned a False Positive AR Vacant Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	110	89.1	5.3	10.9	5.3
Reduced Contacts without AR	117	84.6	8.1	15.4	8.1

Table 5 shows the housing unit type for the resolved cases from Table 4. For resolved housing units in the Control Panel, multiunits accounted for 75.5 percent of false positive AR vacant designations. The analogous statistic for the Reduced Contacts without AR Panel was 69.7 percent. These results show that more differences were observed in multi-unit structures than in single unit.

Table 5: Housing Unit Type for Resolved Cases Assigned a False Positive AR Vacant Flag by Panel

Panel	# Units N	Single Unit		Multi Unit	
		%	SE	%	SE
Control	98	24.5	12.3	75.5	12.3
Reduced Contacts without AR	99	30.3	7.7	69.7	7.7

Conflicting Information Cases

Table 6 shows the number of housing units where, during any one of the interviews associated with that unit, a respondent reported that no one lived there on Census Day. Even though the NRFU status was determined to be occupied, this may call into question the validity of that determination. In the Control Panel, 6.4 percent of housing units had a respondent say, during at least one interview, that no one lived there on Census Day. In the Reduced Contacts without AR Panel, the value was 5.1 percent. These results help explain 5 to 6.4 percent of the differences seen but still leave a large quantity still to explain why administrative records and fieldwork were making different determinations.

Table 6: Did Household Say No One Lived There on Census Day?

Panel	# Units N	Anyone=No	
		%	SE
Control	110	6.4	2.6
Reduced Contacts without AR	117	5.1	2.5

Further examination of the 110 NRFU occupied cases in the Control Panel yielded the following results. In examining the USPS National Change of Address information, we identified that 23 of the 110 had a change within 2 months of July 1, 2014.

Also in examining the 110 cases, we identified that 51 cases had the following:

- 6 field notes indicate vacant
- 8 field notes about moving around July 1st
- 4 field notes indicating possible delete
- 11 interview could not determine population count
- 14 complete by Internet or Telephone Questionnaire Assistance with no self-response vacant option
- 8 building access issues

In Table 7, we show how often the final population count for the false positive AR vacant housing units was resolved. It stands out that for 17.3 percent in the Control Panel and 22.2 percent in the Reduced Contacts without AR Panel, that the interview determined the unit to be occupied, yet the household population was not determined by the interview and remained unresolved. This illustrates another difference between the AR sources and NRFU outcomes, specifically that all AR occupied cases had resolved population counts, whereas not all NRFU enumerations did. These results show that 51 of the 110 cases had differences were observed had some possible explanation when looking further into the notes.

Table 7: Resolved Household Size Status on Census Day for AR Vacant and NRFU Occupied Cases by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	110	82.7	3.6	17.3	3.6
Reduced Contacts without AR	117	77.8	5.9	22.2	5.9

Using information from Tables 6 and 7, 19 percent of the AR vacant units in the Control Panel had some conflicting or incomplete information. This was about half of the cases where there were differences between AR vacant and census occupied determinations.

The subsequent Table 8 shows, for resolved final population cases only, the breakdown of the final population totals. This gives an indication if the AR vacant designations were problematic for certain housing unit sizes. These results make sense intuitively, in that smaller households would more likely be assigned a vacant status than larger households would. This result is expected since smaller households may be more likely to be thought of as vacant.

Table 8: Household Size on Census Day for Resolved AR Vacant and NRFU Occupied Cases by Panel

Panel	# Units N	Pop 1-2		Pop 3-5		Pop 6+	
		%	SE	%	SE	%	SE
Control	91	72.5	8.3	25.3	8.2	2.2	1.6
Reduced Contacts without AR	91	82.4	4.5	14.3	4.2	3.3	1.9

UAA Reason Codes

For the 2014 Census Test, the processing used additional UAA reason codes to flag cases as AR vacant compared to those used in the 2013 Census Test. For more information on the 2013 Census Test, see Walejko et al. (2013). The 2013 Census Test showed that there were NRFU vacant cases not flagged as AR vacant because they had a UAA reason code other than vacant. The question is whether introducing more UAA reason codes to flag cases as AR vacant introduces more potential differences with the census fieldwork results.

For Table 9, the results show the false positive and true positive cases from the Control Panel and the Reduced Contacts without AR Panel. For the Control Panel, there were 110 AR vacant cases that were NRFU occupied, and 61 AR vacant cases that were NRFU non-existent. Analogously, for the Reduced Contacts without AR Panel, these totals were 117 and 101. This results in 389 false positive cases. Similar accounting results in 453 true positive cases.

Table 9 below shows the UAA reason code distribution associated with each category. Specifically, about 32.9 percent of the false positive cases were for UAA reason codes other than vacant. In contrast, only about 19.2 percent of the true positive cases were for UAA reason codes other than vacant. This result shows attempting to identify more cases uses additional UAA reasons than vacant show the potential for more identification but with more differences than seen in the past. UAA vacant continues to be the reason that identifies cases the best.

Table 9: UAA Reason Codes for False Positive and True Positive AR Vacant Cases (Control Panel and the Reduced Contacts without AR Panel Combined)

	# Units N	UAA=Vacant		UAA=Unable to Forward		UAA=Other	
		%	SE	%	SE	%	SE
False Positive	389	67.1	6.8	25.4	6.2	7.5	2.7
True Positive	453	80.8	4.8	16.6	4.4	2.6	1.3

5.1.1.2 False Negative Analysis

This section studied the false negative rate, i.e. the cases that were not assigned a vacant flag, but were vacant in the field. In this analysis, the goal is to determine why the processing did not identify these cases. This can lead us to see if there are ways that we can modify our identifications to account for things that may have been utilized in the production identification.

Vacant Flag Assignment

Table 10 below displays the 3.5 percent and 5.3 percent of cases that were flagged as vacant in the Control Panel and the Reduced Contacts without AR Panel, respectively. The processing did

not assign a vacant flag to 96.5 percent in the Control Panel and 94.7 percent in the Reduced Contacts without AR Panel. This identification was lower than expected.

Table 10: Vacant Flag Assignment by Panel

Panel	# Units N	Vacant Flag Assigned		Vacant Flag NOT Assigned	
		%	SE	%	SE
Control	13,253	3.5	0.6	96.5	0.6
Reduced Contacts without AR	12,553	5.3	0.7	94.7	0.7

Housing Unit Status

Table 11 shows the how often housing units were resolved for cases where a vacant flag was not assigned. The results show that 85.2 percent and 79.1 percent of cases were resolved in the Control Panel and the Reduced Contacts without AR Panel, respectively. These results were lower than expected since the thought was that control panel would have full resolution.

Table 11: Resolved Housing Unit Status for Cases Not Assigned a Vacant Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	12,791	85.2	1.7	14.8	1.7
Reduced Contacts without AR	11,888	79.1	1.9	20.9	1.9

The following table shows what the housing unit status was for these resolved cases. In the Control Panel and the Reduced Contacts without AR Panel respectively, 9.8 percent and 9.7 percent of the cases were vacant.

Table 12: Housing Unit Status for Resolved Cases Not Assigned a Vacant Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	10,904	87.3	1.5	9.8	1.0	2.9	1.0
Reduced Contacts without AR	9,405	88.2	1.2	9.7	1.0	2.1	0.5

Missed Vacant Flag Reasons

Table 13 shows the reasons why a vacant interview was not flagged as vacant. Some of the reasons why a vacant flag was not assigned include:

- the unit was deemed occupied by ARs
- conflicting information – i.e., the unit had information from ARs indicating it was both occupied (see Section 3.1) and vacant (see Section 3.2)
- the unit had an absence of IRS and Medicare data, but the UAA reason code was not defined as being classified as vacant
- the unit had an absence of IRS and Medicare data, no UAA reason code, and the address was missing a secondary number

- the unit had an absence of IRS and Medicare data, no UAA reason code, and the address did not lack any components

In the Control Panel and the Reduced Contacts without AR Panel, respectively, 43.6 percent and 32.1 percent of vacant housing units not assigned a vacant flag were missing a secondary number (i.e., Apt C, or Unit 4). This piece of information may be able to be utilized in the future. The largest percentage was for cases that were assigned occupied based on administrative records. This may be do to trying to use research about determining vacant status on April 1st to a census day of July 1st.

Table 13: Reasons for Missed Vacant Flags for Vacant Interviews by Panel

Panel	# Units	AR Information Present						No AR Information			
		Assigned AR Occupied		Conflicting Information – AR Occupied and AR Vacant		Other UAA Reasons		Address Missing A Secondary Number		No Address Problems	
		%	SE	%	SE	%	SE	%	SE	%	SE
Control	1,069	31.5	2.2	5.2	1.1	2.8	1.2	43.6	5.5	16.8	3.7
Reduced Contacts without AR	910	31.5	2.0	10.9	1.8	4.4	2.1	32.1	4.1	21.1	4.2

5.1.2 Comparing AR Occupied Cases to NRFU Enumerations

This section examines how often the AR assignment of an occupied status matched to the NRFU field status, as well as how often it did not match. To understand the validity of our assignment method, Table 14 first shows the resolved status of housing units, and subsequently, Table 15 shows the final status results of those resolved cases. The resolution results were in line with past resolution results examined.

Table 14: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Control	7,028	89.1	0.9	10.9	0.9
Reduced Contacts without AR	7,002	81.9	1.5	18.1	1.5

Table 15 shows that occupied status assigned by administrative records agreed with the census determination for 93.9 percent and 94.4 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was a good result. Since they were classified as occupied based on administrative records and they were determined to be so in the field, these units can be thought of as true positives. Section 5.1.2.1 will discuss the true positive analysis.

The processing assigned an AR occupied status to a NRFU vacant unit 5.4 percent and 5.0 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. These differences can be thought of as false positives, i.e., the cases for which the processing assigned an occupied status, but were vacant. Section 5.1.2.2 discusses the false negative analysis.

Table 15: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel

Panel	# Units	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	6,259	93.9	0.6	5.4	0.5	0.7	0.3
Reduced Contacts without AR	5,736	94.4	0.7	5.0	0.6	0.6	0.2

5.1.2.1 True Positive Analysis

Household Composition

The following Table 16 shows the NRFU resolved status broken down by the AR household composition. The overall resolved rate in the Control Panel, 89.1 percent, was higher than that of the Reduced Contacts without AR Panel at 81.9 percent. This result was inline with the past resolution results examined. There was no expectation that the control panel would be higher.

Table 16: Resolved Housing Unit Status by Household Composition for AR Occupied Cases by Panel

Panel	HH Composition	# Units	Resolved		Unresolved	
			%	SE	%	SE
Control	1 Adult, 0 Children	2,703	86.2	1.0	13.8	1.0
	1 Adult, 1+ Children	478	88.7	1.7	11.3	1.7
	2 Adults, 0 Children	1,361	89.0	1.3	11.0	1.3
	2 Adults, 1+ Children	920	91.3	1.1	8.7	1.1
	3 Adults, 0 Children	473	93.4	1.4	6.6	1.4
	3 Adults, 1+ Children	349	91.4	1.6	8.6	1.6
	Other	744	93.0	1.3	7.0	1.3
	Total	7,028	89.1	0.9	10.9	0.9
Reduced Contacts without AR	1 Adult, 0 Children	2,732	79.8	2.0	20.2	2.0
	1 Adult, 1+ Children	332	74.1	3.0	25.9	3.0
	2 Adults, 0 Children	1,536	82.8	2.0	17.2	2.0
	2 Adults, 1+ Children	899	83.1	1.4	16.9	1.4
	3 Adults, 0 Children	500	84.4	1.7	15.6	1.7
	3 Adults, 1+ Children	322	90.1	1.5	9.9	1.5
	Other	681	84.9	1.8	15.1	1.8
	Total	7,002	81.9	1.5	18.1	1.5

Table 17 shows, for resolved cases only, the NRFU housing unit status broken down by the AR household composition. In the Control Panel, the results show that ‘1 Adult, 0 Children’, ‘2 Adults, 0 Children’, and ‘2 Adults, 1+ Children’ households comprised 74.6 percent of the true positive occupied cases. These three household categories also matched an occupied housing unit status over 90 percent of the time. These results were similar to the Reduced Contacts without AR Panel. This result was promising that all household compositions had over 90 percent agreement.

Table 17: Housing Unit Status by Household Composition for Resolved AR Occupied Cases by Panel

Panel	HH Composition	# Units	Occupied		Vacant		Non-Existent	
			N	%	SE	%	SE	%
Control	1 Adult, 0 Children	2,331	91.1	0.7	7.7	0.6	1.2	0.5
	1 Adult, 1+ Children	424	96.2	1.1	3.5	1.1	0.2	0.2
	2 Adults, 0 Children	1,211	94.1	1.2	5.4	1.1	0.6	0.3
	2 Adults, 1+ Children	840	96.2	0.8	3.5	0.8	0.4	0.2
	3 Adults, 0 Children	442	98.0	0.7	1.6	0.6	0.5	0.3
	3 Adults, 1+ Children	319	95.9	0.9	3.4	0.9	0.6	0.4
	Other	692	95.2	1.0	4.5	1.0	0.3	0.2
	Total	6,259	93.9	0.6	5.4	0.5	0.7	0.3
Reduced Contacts without AR	1 Adult, 0 Children	2,181	91.7	0.8	7.6	0.8	0.7	0.3
	1 Adult, 1+ Children	246	94.3	1.9	4.5	1.4	1.2	1.2
	2 Adults, 0 Children	1,272	95.5	0.8	4.2	0.8	0.2	0.2
	2 Adults, 1+ Children	747	95.3	0.9	4.0	0.7	0.7	0.3
	3 Adults, 0 Children	422	96.9	0.8	2.1	0.7	0.9	0.5
	3 Adults, 1+ Children	290	96.9	1.1	2.8	1.0	0.3	0.4
	Other	578	97.8	0.7	1.7	0.6	0.5	0.3
	Total	5,736	94.4	0.7	5.0	0.6	0.6	0.2

Comparing AR and NRFU Population Counts

For the 5,876 and 5,414 true positive occupied cases in the Control Panel and the Reduced Contacts without AR Panel, respectively, the following table shows how often the NRFU household population was resolved. In the Control Panel, 92.6 percent of these cases had a NRFU population count. In the Reduced Contacts without AR Panel, 93.6 percent had a NRFU population count. The resolution percentages were inline with previous resolution percentages examined.

Table 18: Resolved Population Count Status for True Positive AR Occupied Cases by Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Control	5,876	92.6	0.7	7.4	0.7
Reduced Contacts without AR	5,414	93.6	0.6	6.4	0.6

Table 19 shows, for cases with resolved household population counts, how often the AR and NRFU household population counts matched. In the Control Panel and the Reduced Contacts without AR Panel, respectively, the match rates were 54.0 percent and 58.3 percent. Since Census Day was July 1st, there was no expectation about how these results should compare. Table 20 shows the resolved status of population counts broken down by the AR household composition.

Table 19: Match Status of Population Counts for Resolved True Positive AR Occupied Cases by Panel

Panel	# Units N	Match		Non-Match	
		%	SE	%	SE
Control	5,442	54.0	1.8	46.0	1.8
Reduced Contacts without AR	5,065	58.3	1.6	41.7	1.6

Table 20: Resolved Population Count Status by Household Composition for True Positive AR Occupied Cases by Panel

Panel	HH Composition	# Units N	Resolved		Unresolved	
			%	SE	%	SE
Control	1 Adult, 0 Children	2,123	92.1	0.9	7.9	0.9
	1 Adult, 1+ Children	408	88.5	1.7	11.5	1.7
	2 Adults, 0 Children	1,139	94.9	1.0	5.1	1.0
	2 Adults, 1+ Children	808	92.1	1.1	7.9	1.1
	3 Adults, 0 Children	433	92.4	1.3	7.6	1.3
	3 Adults, 1+ Children	306	94.1	1.4	5.9	1.4
	Other	659	93.0	1.3	7.0	1.3
	Total	5,876	92.6	0.7	7.4	0.7
Reduced Contacts without AR	1 Adult, 0 Children	2,000	93.7	0.7	6.4	0.7
	1 Adult, 1+ Children	232	87.9	2.3	12.1	2.3
	2 Adults, 0 Children	1,215	94.8	0.7	5.2	0.7
	2 Adults, 1+ Children	712	93.7	0.8	6.3	0.8
	3 Adults, 0 Children	409	93.4	1.3	6.6	1.3
	3 Adults, 1+ Children	281	93.6	1.6	6.4	1.6
	Other	565	92.7	1.5	7.3	1.5
	Total	5,414	93.6	0.6	6.4	0.6

For resolved population count cases only, Table 21 shows the population count match rates for various household compositions. In the Control Panel, the results show that ‘1 Adult, 0 Children’, ‘2 Adult, 0 Children’, and ‘2 Adult, 1+ Children’ households comprised 69.5 percent of the matched population count cases for true positive occupied cases. These three household compositions all had match rates over 50 percent, which was significantly higher than any of the other household compositions. These results were similar to the Reduced Contacts without AR Panel. The better performance for these three composition categories was one of the major findings of our test. This was something not known before.

Table 21: Match Status of Population Counts by Household Composition for Resolved True Positive AR Occupied Cases by Panel

Panel	HH Composition	# Units N	Match		Non-Match	
			%	SE	%	SE
Control	1 Adult, 0 Children	1,955	65.3	2.7	34.7	2.7
	1 Adult, 1+ Children	361	40.4	2.6	59.6	2.6
	2 Adults, 0 Children	1,081	61.4	1.5	38.6	1.5
	2 Adults, 1+ Children	744	56.6	2.6	43.4	2.6
	3 Adults, 0 Children	400	33.3	2.5	66.8	2.5
	3 Adults, 1+ Children	288	35.4	3.1	64.6	3.1
	Other	613	32.3	2.0	67.7	2.0
	Total	5,442	54.0	1.8	46.0	1.8
Reduced Contacts without AR	1 Adult, 0 Children	1,873	69.2	2.8	30.8	2.8
	1 Adult, 1+ Children	204	37.3	2.9	62.7	2.9
	2 Adults, 0 Children	1,152	65.5	1.5	34.5	1.5
	2 Adults, 1+ Children	667	62.1	2.8	37.9	2.8
	3 Adults, 0 Children	382	44.0	2.4	56.0	2.4
	3 Adults, 1+ Children	263	39.5	3.1	60.5	3.1
	Other	524	27.1	1.8	72.9	1.8
	Total	5,065	58.3	1.6	41.7	1.6

For the cases that did not match, the next analysis done examined how often the NRFU population count exceeded the AR population count, and vice versa. The results are shown in Table 22. In the Control Panel, 56.6 percent of cases had a population count greater in NRFU than in AR. The analogous percentage in the Reduced Contacts without AR Panel was 54.1 percent. This result shows that improvements are still needed for the administrative records count determines since on average they were lower than census responses.

Table 22: Population Count Comparison for Resolved, Non-Matched True Positive AR Occupied Cases by Panel

Panel	# Units N	Greater in NRFU		Greater in AR	
		%	SE	%	SE
Control	2,502	56.6	1.6	43.4	1.6
Reduced Contacts without AR	2,110	54.1	1.7	45.9	1.7

Table 23 shows the NRFU vs. AR population count comparison broken down by the AR household composition. In both the Control Panel and the Reduced Contacts without AR Panel, the first three household compositions listed, which are generally smaller households, have over 50 percent of their cases “Greater in NRFU”. The subsequent four household compositions, generally larger households, have over 50 percent of their cases “Greater in AR.” This makes sense intuitively given the household composition is based on AR data.

Table 23: Population Count Comparison by Household Composition for Resolved, Non-Matched True Positive AR Occupied Cases by Panel

Panel	HH Composition	# Units N	Greater in NRFU		Greater in AR	
			%	SE	%	SE
Control	1 Adult, 0 Children	679	100.0	0.0		
	1 Adult, 1+ Children	215	64.7	3.6	35.3	3.6
	2 Adults, 0 Children	417	55.2	3.7	44.8	3.7
	2 Adults, 1+ Children	323	44.6	3.8	55.4	3.8
	3 Adults, 0 Children	267	23.6	3.0	76.4	3.0
	3 Adults, 1+ Children	186	16.7	2.4	83.3	2.4
	Other	415	31.1	5.0	68.9	5.0
	Total	2,502	56.6	1.6	43.4	1.6
Reduced Contacts without AR	1 Adult, 0 Children	576	100.0	0.0		
	1 Adult, 1+ Children	128	51.6	5.3	48.4	5.3
	2 Adults, 0 Children	398	55.5	3.8	44.5	3.8
	2 Adults, 1+ Children	253	41.1	3.8	58.9	3.8
	3 Adults, 0 Children	214	28.0	3.0	72.0	3.0
	3 Adults, 1+ Children	159	21.4	4.2	78.6	4.2
	Other	382	20.9	2.2	79.1	2.2
	Total	2,110	54.1	1.7	45.9	1.7

The following table shows another version of the AR and NRFU Population Count Comparison based on comparison of population count differences. In the Control Panel, the population count was within one person for 82.8 percent of the cases. In the Reduced Contacts without AR Panel, this value was 84.1 percent.

Table 24: Population Count Comparison for Resolved True Positive AR Occupied Cases by Panel

Panel	# Units N	2 or More Fewer Persons		1 Fewer Person		Match		1 Greater Person		2 or More Greater Persons	
		%	SE	%	SE	%	SE	%	SE	%	SE
Control	5,442	7.1	0.8	12.9	0.6	54.0	1.8	15.9	0.8	10.1	0.8
Reduced Contacts without AR	5,065	7.9	0.8	11.2	0.6	58.3	1.6	14.6	0.7	8.0	0.6

Note: Table 24 should read as NRFU count relative to AR count. So “1 Fewer Person” means, for example, that NRFU count=2 and AR count=3.

Table 25 documents the percent difference in the NRFU population count and the AR population count when aggregated to the site level. The total site aggregate (DC+MD) is also presented in the table.

Table 25: Comparison of AR and NRFU Population Counts for Resolved Cases by Panel

Panel	State	NRFU Pop	AR Pop	Total Pop Difference	% Difference*
		N	N	N	%
Control	DC	3,190	2,951	239	7.5%
	MD	10,734	10,254	480	4.5%
	Total	13,924	13,205	719	5.2%
Reduced Contacts without AR	DC	3,515	3,394	121	3.4%
	MD	8,801	8,687	114	1.3%
	Total	12,316	12,081	235	1.9%

*% Difference = ((NRFU Pop-AR Pop) / NRFU Pop)*100

The tables in this section document the comparisons of Administrative record occupied cases to census field. The overall results show over 90 percent agreement in occupancy status. The results showing only about 54 percent agreement in counts and instances where census fieldwork is producing larger counts are areas to continue to address. One possibility is reducing the amount of cases identified to try to improve the comparison results seen here.

5.1.2.2 False Positive Analysis

For the 624 false positives in the Control Panel and the Reduced Contacts without AR Panel, this section documents why a case determined to be occupied based on administrative records was determined to be vacant in the field.

Housing Unit Type

Table 26 shows the resolved housing unit type status for the false positive cases assigned a vacant status in the field. In the Control Panel, 93.2 percent of the cases had a resolved housing unit type status, while the Reduced Contacts without AR Panel cases were resolved 92.7 percent of the time. These resolution percentages were in line with the previous resolution percentages examined.

Table 26: Resolved Housing Unit Type Status for Cases Assigned a False Positive AR Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	337	93.2	3.3	6.8	3.3
Reduced Contacts without AR	287	92.7	4.5	7.3	4.5

Table 27 shows the housing unit type, for resolved cases only, for the false positive cases assigned a NRFU vacant status. For resolved housing units in the Control Panel, multi-units accounted for 68.5 percent of false positive occupied designations determined to be vacant in NRFU. The analogous statistic for the Reduced Contacts without AR Panel was 67.7 percent.

These results document that more differences were observed in multi-unit areas. Further work should look to see what enhancements can be done for addresses in those types of structures.

Table 27: Housing Unit Type for Resolved Cases Assigned a False Positive AR Occupied Flag by Panel

Panel	# Units N	Single Unit		Multi Unit	
		%	SE	%	SE
Control	314	31.5	7.0	68.5	7.0
Reduced Contacts without AR	266	32.3	6.8	67.7	6.8

Conflicting Information Cases

Table 28 shows the number of housing units where, during any one of the interviews associated that unit, a respondent (proxy) reported that someone lived there on Census Day (Anyone = Yes). Even though the NRFU status was vacant, this may call into question the validity of that determination.

In the Control Panel, 3.0 percent of housing units had a respondent say, during at least one interview, that someone lived there on Census Day. In the Reduced Contacts without AR Panel, this value was 2.4 percent. This turned out to be a very small part of the difference.

Table 28: Did Household Say Anyone Lived there on Census Day?

Panel	# Units N	Anyone=Yes	
		%	SE
Control	337	3.0	0.9
Reduced Contacts without AR	287	2.4	1.0

5.1.2.3 False Negative Analysis

This section documents the examination of the the false negative rate, i.e., the cases that were not assigned an occupied flag based on administrative records, but were occupied in the field.

Occupied Flag Assignment

Table 29 below shows that of the 13,253 cases in the Control Panel, 47.0 percent were not assigned an occupied flag. In the Reduced Contacts without AR Panel, 44.2 percent were not assigned an occupied flag. This was smaller than past research in using 2010 data where approaches had around 70 percent of the cases not being assigned occupied.

Table 29: Occupied Flag Assignment By Panel

Panel	# Units N	Occupied Flag Assigned		Occupied Flag NOT Assigned	
		%	SE	%	SE
Control	13,253	53.0	3.1	47.0	3.1
Reduced Contacts without AR	12,553	55.8	1.6	44.2	1.6

Housing Unit Status

Of the 6,225 cases not assigned an occupied flag in the Control Panel, 80.5 percent had a resolved housing unit status. In the Reduced Contacts without AR Panel, of the 5,551 cases not assigned an occupied flag, 74.7 percent were resolved. For the control panel, the resolution was lower than expected. Since the reduced contacts had less fieldwork, that result was not projected.

Table 30: Resolved Housing Unit Status for Cases Not Assigned an Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	6,225	80.5	2.2	19.5	2.2
Reduced Contacts without AR	5,551	74.7	2.5	25.3	2.5

Table 31 shows that 74.8 percent of resolved cases not assigned an occupied flag were determined to be occupied in NRFU. In the Reduced Contacts without AR Panel, this analogous statistic is 72.4 percent. This result does show that there is a large magnitude of occupied cases remaining that potentially could be identified. The remaining parts of this section examines if they are associated with particular characteristics. If so, that could be utilized in future work to help improve identification.

Table 31: Housing Unit Status for Resolved Cases Not Assigned an Occupied Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	5,011	74.8	1.6	18.5	1.4	6.7	1.6
Reduced Contacts without AR	4,145	72.4	2.2	21.3	2.0	6.4	1.5

Missed Occupied Flag Reasons

The reasons these cases were not identified as occupied before they reached the field include:

- assigned AR vacant
- conflicting information – i.e., the unit had information from ARs indicating it was both occupied (see Section 3.1) and vacant (see Section 3.2)
- the ARs indicated a population greater than six
- the unit had an absence of IRS and Medicare data

Table 32 below shows the reasons why an occupied interview was not flagged as occupied. The largest reason, at 87.5 percent and 87.0 percent in the Control Panel and the Reduced Contacts without AR Panel, respectively, was that the addresses were missing AR information from IRS, Medicare or both. The main reason was the absence of IRS and/or Medicare information. This result was expected to be the largest cause.

Table 32: Reasons for Missed Occupied Flags for Occupied Interviews by Panel

Panel	# Units	AR Information Present			No AR Information
		Assigned AR Vacant	Conflicting Information – AR Occupied and AR Vacant	AR Pop Greater than 6	

	N	%	SE	%	SE	%	SE	%	SE
Control	3,749	2.9	0.6	1.8	0.3	7.8	1.8	87.5	2.0
Reduced Contacts without AR	2,999	3.9	0.5	2.5	0.4	6.6	1.1	87.0	1.2

5.2 Reduced Contacts with Full AR Removal Panel Analysis

For the Reduced Contacts with Full AR Removal Panel, the processing removed the AR occupied and AR vacant cases before going into the field. As a result, this section analyzes and attempted to increase our understanding of the occupied and vacant cases resolved after going into the field.

Housing Unit Status

To begin, of the 8,101 cases in the Reduced Contacts with Full AR Removal Panel, only 389 were called AR vacant and 4,656 were called administrative record occupied. These cases were removed from the NRFU workload and the remaining 3,056 were sent into the field. Of these cases sent to the field, 72.1 percent had a resolved housing unit status. Since this panel had reduced fieldwork, there was no expectation of what the result would be.

Table 33: Resolved Housing Unit Status of Cases Not Removed in the Reduced Contacts with Full AR Removal Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Reduced Contacts with Full AR Removal	3,056	72.1	2.6	27.9	2.6

The housing unit statuses of the resolved cases are shown in Table 34 below. In the Reduced Contacts with Full AR Removal Panel, 79.2 percent of resolved cases were occupied, 17.7 percent were vacant, and 3.1 percent were non-existent. This result was as expected that most of the cases remaining were occupied.

Table 34: Housing Unit Status of Resolved Cases Not Removed in the Reduced Contacts with Full AR Removal Panel

Panel	# Units	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Reduced Contacts with Full AR Removal	2,203	79.2	2.5	17.7	2.4	3.1	0.9

Missed Vacant Flag Reasons

For the 390 vacant designation cases, Table 35 shows the reasons they were not assigned as vacant based on administrative records before they reached the field. These were the same reasons for the false negatives in the Control Panel and the Reduced Contacts without AR Panel. The main reason was that no AR information including UAA reasons were available for the unit. In examining further in the Reduced Contacts with Full AR Removal Panel, 63.1 percent of missed vacant flags were missing a secondary number, i.e., unit designation, such as, Apt 401.

The UAA reasons not being available was expected and the missing secondary number was a new finding.

Table 35: Reasons for Missed Vacant Flags for Vacant Interviews for the Reduced Contacts with Full AR Removal Panel

Panel	# Units	AR Information Present						No AR Information			
		Assigned AR Occupied		Conflicting Information – AR Occupied and AR Vacant		Other UAA Reasons		Address Missing A Secondary Number		No Address Problems	
	N	%	SE	%	SE	%	SE	%	SE	%	SE
Reduced Contacts with Full AR Removal	390	0.0	0.0	9.7	3.0	3.8	1.8	63.1	11.2	23.3	8.0

Missed Occupied Flag Reasons

In the Reduced Contacts with Full AR Removal Panel, 79.2 percent of resolved cases were determined to be occupied. Table 36 shows the reasons why an occupied interview was not flagged as occupied. The largest reason was that the addresses were missing AR information from IRS, Medicare or both. This result matched expectations based on previous research of 2010 data.

Table 36: Reasons for Missed Occupied Flags for Occupied Interviews for the Reduced Contacts with Full AR Removal Panel

Panel	# Units	AR Information Present						No AR Information	
		Assigned AR Vacant		Conflicting Information – AR Occupied and AR Vacant		AR Pop Greater than 6			
	N	%	SE	%	SE	%	SE	%	SE
Reduced Contacts with Full AR Removal	1,744	0.0	0.0	1.3	0.3	8.2	2.0	90.5	2.1

5.3 Adaptive Design with Hybrid AR Removal Panel Analysis

Of the 12,340 cases in the Adaptive Design with Hybrid AR Removal Panel, 507 were called AR vacant. These cases were flagged and removed from the NRFU workload before going into the field. In addition, 6,926 were flagged as AR occupied. One visit was allowed to these cases. Of these 6,926 cases, 3,420 were resolved on the first contact. The remaining 3,506 AR occupied cases were not resolved on the first visit. These received no more contacts during the NRFU operation and were enumerated using ARs. This section examines the comparison of administrative record occupied results and the census results from the first contact.

Housing Unit Status

To begin, there were 3,420 cases that were flagged as AR occupied, but were resolved on this first contact. By definition, all of these cases should be resolved. However, there were a few cases, specifically 1.2 percent, that were unresolved and were erroneously pulled as being resolved on the first contact.

Table 37: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag in the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Adaptive Design with Hybrid AR Removal	3,420	98.8	0.2	1.2	0.2

As shown in Table 38, the occupied status agreed with the NRFU status 97.5 percent of the time. Since these cases were classified as occupied based on administrative records and they were determined to be so in the field, these units can be thought of as true positives. It was a good result to see that this percentage was as high as seen.

Table 38: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag in the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Adaptive Design with Hybrid AR Removal	3,378	97.5	0.4	2.0	0.3	0.5	0.1

Household Composition

The next Table 39 shows the breakdown of resolved status by AR household compositions. The resolved rates were similar for the different household composition groups. These resolved rates, again, in theory, should all be 100 percent. The only reason for results less than 100 percent were described above.

Table 39: Resolved Housing Unit Status by Household Composition for AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel

Panel	HH Composition	# Units N	Resolved		Unresolved	
			%	SE	%	SE
Adaptive Design with Hybrid AR Removal	1 Adult, 0 Children	1,084	98.1	0.4	1.9	0.4
	1 Adult, 1+ Children	204	98.0	1.1	2.0	1.1
	2 Adults, 0 Children	772	98.7	0.4	1.3	0.4
	2 Adults, 1+ Children	474	99.6	0.3	0.4	0.3
	3 Adults, 0 Children	275	99.3	0.5	0.7	0.5
	3 Adults, 1+ Children	240	99.6	0.4	0.4	0.4
	Other	371	99.5	0.4	0.5	0.4
Total	3,420	98.8	0.2	1.2	0.2	

Table 40 shows the household composition breakdown for each housing unit status. Each category had a high rate of being occupied. The ‘1 Adult, 0 Children’ category had a vacant rate of 3.5 percent. One possible explanation was that this was the smallest household composition size and thus making it possibly more likely to have a vacant designation. This was a very good

result to see that units matched 95 percent or more over the different household composition groups.

Table 40: Housing Unit Status by Household Composition for Resolved AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel

Panel	HH Composition	# Units	Occupied		Vacant		Non-Existent	
			%	SE	%	SE	%	SE
Adaptive Design with Hybrid AR Removal		N						
	1 Adult, 0 Children	1,063	95.6	0.7	3.5	0.6	0.9	0.3
	1 Adult, 1+ Children	200	98.0	0.9	1.5	0.7	0.5	0.5
	2 Adults, 0 Children	762	97.4	0.6	2.2	0.5	0.4	0.2
	2 Adults, 1+ Children	472	99.4	0.4	0.4	0.3	0.2	0.2
	3 Adults, 0 Children	273	99.3	0.5	0.0	0.0	0.7	0.5
	3 Adults, 1+ Children	239	99.2	0.6	0.8	0.6	0.0	0.0
	Other	369	98.4	0.8	1.4	0.6	0.3	0.3
Total	3,378	97.5	0.4	2.0	0.3	0.5	0.1	

Comparing AR and NRFU Population Counts

For the 3,294 true positive occupied cases in the Adaptive Design with Hybrid AR Removal Panel, all but one of the cases had a resolved population count. Due to rounding, however, the table shows 100.0 percent resolved cases. Table 41 shows these results.

Table 41: Resolved Population Count Status for True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Adaptive Design with Hybrid AR Removal	3,294	100.0	0.0	0.0	0.0

The following Table 42 shows how often the AR and NRFU household population counts matched for resolved cases. For cases with resolved household population counts, 56.7 percent of cases matched AR and NRFU population counts in the Adaptive Design with Hybrid AR Removal Panel. Table 43 shows the one unresolved case came from the ‘1 Adult, 1+ Children’ category. This result was similar to the comparisons shown earlier for the control panel. Agreement of only 56.7 percent does lead us to consider other ways to improve the identification. These responses were from household members so this does suggest benefits from conducting visits before using administrative records for occupied units.

Table 42: Match Status of Population Counts for Resolved True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units	Match		Non-Match	
		%	SE	%	SE
Adaptive Design with Hybrid AR Removal	3,293	56.7	1.2	43.3	1.2

Table 43: Resolved Population Count Status by Household Composition for True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel

Panel	HH Composition	# Units N	Resolved		Unresolved	
			%	SE	%	SE
Adaptive Design with Hybrid AR Removal	1 Adult, 0 Children	1,016	100.0	0.0	0.0	0.0
	1 Adult, 1+ Children	196	99.5	0.5	0.5	0.5
	2 Adults, 0 Children	742	100.0	0.0	0.0	0.0
	2 Adults, 1+ Children	469	100.0	0.0	0.0	0.0
	3 Adults, 0 Children	271	100.0	0.0	0.0	0.0
	3 Adults, 1+ Children	237	100.0	0.0	0.0	0.0
	Other	363	100.0	0.0	0.0	0.0
	Total	3,294	100.0	0.0	0.0	0.0

Table 44 shows the match rates for various household compositions. The results show that, similar to the Control Panel and the Reduced Contacts without AR Panel, that ‘1 Adult, 0 Children’, ‘2 Adults, 0 Children’, and ‘2 Adults, 1+ Children’ household compositions performed well when matching AR and NRFU population counts. These three household compositions had match rates of 66.1 percent, 63.9 percent, and 66.7 percent, respectively. One of our new findings is better count agreements of these categories. It is good to see that the same relationship seen earlier when looking at the control panel is present for these household respondents who had one contact.

Table 44: Match Status of Population Counts by Household Composition for Resolved True Positive AR Occupied Cases in the Adaptive Design with Hybrid AR Removal Panel

Panel	HH Composition	# Units N	Match		Non-Match	
			%	SE	%	SE
Adaptive Design with Hybrid AR Removal	1 Adult, 0 Children	1,016	66.1	2.8	33.9	2.8
	1 Adult, 1+ Children	195	37.9	3.1	62.1	3.1
	2 Adults, 0 Children	742	63.9	1.8	36.1	1.8
	2 Adults, 1+ Children	469	66.7	2.9	33.3	2.9
	3 Adults, 0 Children	271	38.7	2.9	61.3	2.9
	3 Adults, 1+ Children	237	44.7	3.2	55.3	3.2
	Other	363	34.2	2.0	65.8	2.0
	Total	3,293	56.7	1.2	43.3	1.2

For the cases that did not match, the results show how often the NRFU population count exceeded the AR population count, and vice versa. These results are shown in Table 45, where for 54.9 percent of the time, NRFU enumerations yielded higher population counts than AR sources. This result is a concern since administrative records may be missing some of the population.

Table 45: Population Count Comparison for Resolved, Non-Matched True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units N	Greater in NRFU		Greater in AR	
		%	SE	%	SE
Adaptive Design with Hybrid AR Removal	1,425	54.9	1.6	45.1	1.6

The next Table 46 shows the non-matched population count comparison broken down by household compositions. The general trend was that as household populations increased in size, according to the AR sources, the more likely that population count would be greater in AR relative to NRFU. Also, note that since household compositions were based on the AR sources, and by definition, a ‘1 Adult, 0 Children’ designation was a household population of one, that it would be impossible to have any ‘Greater in AR’ cases for NRFU occupied housing units in this category. Looking at composition was a new point of research so this was a new finding that we were able to learn. It was similar to past results for results seen based on administrative record counts.

Table 46: Population Count Comparison by Household Composition for Resolved, Non-Matched True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel

Panel	HH Composition	# Units	Greater in NRFU		Greater in AR	
			%	SE	%	SE
Adaptive Design with Hybrid AR Removal	1 Adult, 0 Children	344	100.0	0.0	0.0	0.0
	1 Adult, 1+ Children	121	67.8	5.2	32.2	5.2
	2 Adults, 0 Children	268	56.7	4.9	43.3	4.9
	2 Adults, 1+ Children	156	43.6	4.7	56.4	4.7
	3 Adults, 0 Children	166	31.3	3.0	68.7	3.0
	3 Adults, 1+ Children	131	21.4	3.0	78.6	3.0
	Other	239	23.4	3.0	76.6	3.0
	Total	1,425	54.9	1.6	45.1	1.6

Table 47 shows another version of the AR and NRFU population count comparison. This version shows how far off the AR and NRFU sources are from one another. In the Adaptive Design with Hybrid AR Removal Panel, 83.4 percent of the cases were matched or within one person. This result is good and is in line with past research.

Table 47: Population Count Comparison for Resolved True Positive AR Occupied Cases for the Adaptive Design with Hybrid AR Removal Panel

Panel	# Units	2 or More Fewer Persons		1 Fewer Person		Match		1 Greater Person		2 or More Greater Persons	
		%	SE	%	SE	%	SE	%	SE	%	SE
Adaptive Design with Hybrid AR Removal	3,293	6.9	0.6	12.6	0.7	56.7	1.2	14.0	0.7	9.7	0.7

Note: Table 47 should read as NRFU count relative to AR count. So “1 Fewer Person” means, for example, that NRFU count=2 and AR count=3.

5.4 ‘What If?’ Scenarios

Because the Control Panel and the Reduced Contacts without AR Panel completed interviews without any use of ARs, this section documents how we altered the methodologies to flag occupied and vacant cases, and then completed the same type of NRFU comparison analysis completed in Section 5.1. This allows us to see how different usages might affect identification and quality.

5.4.1 What if only the UAA Vacant reason code was used to flag cases as vacant?

To assign an AR vacant flag in the 2014 Census Test, the assignment used the presence of multiple UAA reason codes other than vacant. These codes are described in Section 3. In the 2013 Census Test, only the UAA reason code of vacant was used to flag cases as vacant. This scenario shows the results of using this 2013 approach. The remaining conditions for flagging AR occupied cases are consistent with 2014 methods. In doing so, 58.7 percent of cases, across all panels, would have been identified as AR occupied or AR vacant. In the 2014 production scenario, this value was 59.8 percent.

Table 48 provides the distribution of initial flags for the Control Panel and the Reduced Contacts without AR Panel. It sets up the universe of AR vacant and AR occupied cases to be analyzed in the ensuing tables.

Table 48: AR Initial Flags by Panel: Alternate AR Removal Scenario Allowing Only a UAA Vacant Reason Code for Flagging AR Vacant Removals

Panel	# Units	Initial Flags					
		No AR Identification		AR Occupied		AR Vacant	
	N	%	SE	%	SE	%	SE
Control	13,253	44.4	3.3	53.4	3.2	2.2	0.6
Reduced Contacts without AR	12,553	39.4	1.5	56.2	1.6	4.4	0.7

This part focuses on how often our AR assignment of a vacant status matched to the NRFU field status, as well as how often it did not match. Table 49 shows the percentage of resolved housing unit statuses for the 286 and 548 cases in the Control Panel and the Reduced Contacts without AR Panel, respectively. Note that these values were less than the number of vacant cases that were removed in the production scenario, because this approach used fewer UAA reason codes to assign a vacant flag.

Table 49: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Control	286	78.3	3.3	21.7	3.3
Reduced Contacts without AR	548	73.5	4.6	26.5	4.6

To understand the validity of our assignment method, Table 50 shows the housing unit status of resolved AR vacant cases in the Control Panel and the Reduced Contacts without AR Panel. Using the rules in Scenario 5.4.1, this approach would have agreed on a vacant status 62.1 percent and 56.3 percent of the time, and an unoccupied status 77.7 percent and 78.9 percent of the time, in the Control Panel and the Reduced Contacts without AR Panel, respectively. Lastly, this approach would have assigned an AR vacant status to an NRFU occupied unit 22.3 percent and 21.1 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively.

Table 50: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	224	22.3	3.3	62.1	6.0	15.6	8.0
Reduced Contacts without AR	403	21.1	3.6	56.3	8.8	22.6	9.9

This part examined how often our AR assignment of an occupied status matched to the NRFU field status, as well as how often it did not match. Notice that for the Control Panel and the Reduced Contacts without AR Panel, respectively, 7,081 and 7,053 cases were analyzed in the Table 51, which was slightly higher than the production number. This was because there were fewer cases with AR vacant flags due to the use of fewer UAA reason codes, and thus there were fewer conflicting information cases (those with both an AR vacant and an AR occupied flag). It follows that some cases designated as conflicting information cases in production were just AR occupied cases in this scenario.

Table 51 shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 51: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	7,081	88.9	0.9	11.1	0.9
Reduced Contacts without AR	7,053	81.8	1.6	18.2	1.6

Using the rules in Scenario 5.4.1, this result would have agreed on an occupied status 93.7 percent and 94.2 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about the same as the true positive rate seen in the production scenario. This approach assigned an AR occupied status to an NRFU unoccupied unit 6.3 percent and 5.8 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. These differences can be thought of as false positives, i.e., the cases for that were assigned an occupied status based on administrative record information, but were unoccupied. This was about the same as the false positive rate seen in the production scenario.

Table 52: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	6,294	93.7	0.6	5.5	0.5	0.7	0.3
Reduced Contacts without AR	5,769	94.2	0.7	5.2	0.6	0.6	0.2

This alternate identification scenario would have had the following implications across the Control Panel and the Reduced Contacts without AR combined:

- 1.2 percent reduction in the total number of cases identified based on AR.
- 26.0 percent reduction in the number of cases flagged as AR vacant.
- 0.7 percent increase in the number of cases flagged as AR occupied.

- Under the full AR removal approach, the overall reduction in flagged cases would have increased the NRFU workload by 1.8 percent.
- A 58.4 percent true positive vacant rate. In production, it was 53.8 percent
- A 93.9 percent true positive occupied rate. In production, it was 94.1 percent.

The good news is that this section documents that the results show an increase from 53.8 percent true positive vacant rate to 58.4 percent if we decreased our vacant identification by 26 percent. This is one possible cost trade off that could be done to improve quality based on this rule-based implementation. This change has no impact on administrative record occupied identification.

The following two tables show the levels used to calculate the percent change metrics just mentioned. They also show the results for the Control Panel and the Reduced Contacts without AR separately.

Table 53: Percent Change in AR Removals by Panel

Panel	AR Occupied		% Change	AR Vacant		% Change	Total AR Removals		% Change
	Production	Alt Scenario		Production	Alt Scenario		Production	Alt Scenario	
Control	7,028	7,081	0.8%	462	286	38.1%	7,490	7,367	1.6%
Reduced Contacts without AR	7,002	7,053	0.7%	665	548	17.6%	7,667	7,601	0.9%
Total	14,030	14,134	0.7%	1,127	834	26.0%	15,157	14,968	1.2%

Table 54: Percent Change in NRFU Workload by Panel

Panel	NRFU Universe	Total AR Removals		NRFU Workload*		% Change
		Production	Alt Scenario	Production	Alt Scenario	
Control	13,253	7,490	7,367	5,763	5,886	2.1%
Reduced Contacts without AR	12,553	7,667	7,601	4,886	4,952	1.4%
Total	25,806	15,157	14,968	10,649	10,838	1.8%

*NRFU Workload = NRFU Universe – Total AR Removals. This assumes a full AR removal approach.

5.4.2 What if we only used certain AR Household Compositions to Flag Cases as AR occupied?

To assign an AR occupied flag, the processing did not distinguish between the types of AR households. However, results from Section 5 show better household size agreement between AR and NRFU if the processing might only consider three AR household compositions. The three compositions are ‘1 Adult, 0 Children’, ‘2 Adults, 0 Children’, and ‘2 Adults, 1+ Children.’ This scenario used only these three AR household compositions to flag NRFU units as AR occupied. The remaining conditions for flagging AR vacant cases are consistent with 2014 rules. In doing so, 44.3 percent of cases across all panels would have been identified, compared to 59.8 percent in the production scenario. Table 55 provides the distribution of initial flags for the Control Panel and the Reduced Contacts without AR Panel.

Table 55: AR Initial Flags By Panel: Alternate AR Removal Scenario Allowing Only Specific Household Compositions for Flagging AR Occupied Removals.

Panel	# Units N	No AR Removal		AR Occupied		AR Vacant	
		%	SE	%	SE	%	SE
Control	13,253	58.9	2.0	37.6	1.8	3.5	0.6
Reduced Contacts without AR	12,553	53.5	1.1	41.2	0.9	5.3	0.7

For this part, the processing did not change any procedures for assigning AR vacant cases. Still, there was the possibility that the number of AR vacant cases could change under this alternative scenario. This could have occurred if the following happened. A production case could have had conflicting information where the production rules identified it as both AR occupied and AR vacant. If using this alternative approach, the address could have no longer had its AR occupied flag under the stricter rules of this alternative approach. This would have left the case with just an AR vacant flag, and thus increases in these flags were possible despite no rules directly affecting them. However, this decision was that these cases should remain conflicting information cases. Therefore, the AR vacant results are no different from those seen in production, and all of the analysis in this section revolves around changes in the AR occupied designations.

This part examines how often our AR assignment of an occupied status matched to the NRFU field status, as well as how often it did not match. Notice that there were 4,984 and 5,167 cases respectively analyzed in this table. This reduction relative to production levels was because restrictions were on the AR household composition. Table 56 shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 56: Resolved Housing Unit Status for Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	4,984	87.9	1.0	12.1	1.0
Reduced Contacts without AR	5,167	81.3	1.7	18.7	1.7

Using the rules in Scenario 5.4.2, the results show agreement on occupied status 92.9 percent and 93.5 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about a 1 percent smaller true positive rate than seen in the production scenario. This approach assigned an AR occupied status to an NRFU unoccupied unit 7.1 percent and 6.5 percent of the time in the Control Panel and the Reduced Contacts without AR Panel.

Table 57: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	4,382	92.9	0.7	6.2	0.6	0.9	0.4
Reduced Contacts without AR	4,200	93.5	0.7	5.9	0.7	0.6	0.2

For true positive occupied cases with resolved household population counts, 62.5 percent and 66.8 percent matched AR and NRFU population counts in the Control Panel and the Reduced

Contacts without AR Panel, respectively. In both panels combined, this rate was 64.6 percent, which represents about a 10-percentage points higher match rate from the production methodology.

This alternate identification scenario would have had the following implications across the Control Panel and the Reduced Contacts without AR combined:

- 25.6 percent reduction in the total number of cases identified based on AR
- 27.6 percent reduction in the number of cases flagged as AR occupied
- Under the full AR removal approach, the overall reduction in flagged cases would have increased the NRFU workload by 36.4 percent.
- A 93.2 percent true positive occupied rate. In production, it was 94.1 percent.
- A 64.6 percent population match rate between AR and NRFU sources for true positive occupied cases with resolved population counts. In production, this value was 56.1 percent.

This section shows that the percent count agreement can be increased from 56.1 percent to 64.6 percent if stricter rules based on household compositions are used. While the count agreement increases, it does result in 27.6 percent fewer cases identified by this rule-based approach.

The following two tables show the levels used to calculate the percent change metrics just mentioned. They also show the results for the Control Panel and the Reduced Contacts without AR separately.

Table 58: Percent Change in AR Removals by Panel

Panel	AR Occupied		% Change	AR Vacant		% Change	Total AR Removals		% Change
	Production	Alt Scenario		Production	Alt Scenario		Production	Alt Scenario	
Control	7,028	4,984	29.1%	462	462	0.0%	7,490	5,446	27.3%
Reduced Contacts without AR	7,002	5,167	26.2%	665	665	0.0%	7,667	5,832	23.9%
Total	14,030	10,151	27.6%	1,127	1,127	0.0%	15,157	11,278	25.6%

Table 59: Percent Change in NRFU Workload by Panel

Panel	NRFU Universe	Total AR Removals		NRFU Workload*		% Change
		Production	Alt Scenario	Production	Alt Scenario	
Control	13,253	7,490	5,446	5,763	7,807	35.5%
Reduced Contacts without AR	12,553	7,667	5,832	4,886	6,721	37.6%
Total	25,806	15,157	11,278	10,649	14,528	36.4%

*NRFU Workload = NRFU Universe – Total AR Removals. This assumes a full AR removal approach.

5.4.3 What if match of IRS households between 2012 and 2013 data was required?

The 2014 Census Test used the presence of the 2013 IRS 1040, the 2012 and 2013 Medicare data, and the absence of certain UAA reason codes to assign cases as occupied. This approach resulted in 59.8 percent of the total NRFU universe being identified. In the 2013 Census Test, two years of matching IRS data were used. This was a more conservative approach to identifying NRFU addresses as being good AR occupied units. Applying that construct to the 2014 Census Test would have resulted in only 39.7 percent of cases being identified. Table 60 provides the distribution of initial flags for the Control Panel and the Reduced Contacts without AR Panel.

Table 60: AR Flags By Panel: Alternate AR Removal Scenario Requiring a Match Between IRS 2012 and the 2013 Households for Flagging AR Occupied Removals

Panel	# Units	Initial Flags					
		No AR Identification		AR Occupied		AR Vacant	
		%	SE	%	SE	%	SE
Control	13,253	64.2	2.7	31.9	2.5	3.9	0.6
Reduced Contacts without AR	12,553	58.1	1.5	36.0	1.6	5.9	0.7

This part shows how often our AR assignment of a vacant status matched to the NRFU field status, as well as how often it did not match. Notice that there are 512 and 744 cases respectively analyzed in this table. This was more than the number of vacant cases that were removed in production, because the additional restriction that the 2013 IRS 1040 household must match the 2012 IRS 1040 household was placed. If the households did not match, and there was an eligible UAA code identified in Section 3.2, then the unit was flagged as vacant. Table 61 shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 61: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel

Panel	# Units	Resolved		Unresolved	
		%	SE	%	SE
Control	512	77.7	2.1	22.3	2.1
Reduced Contacts without AR	744	71.2	4.1	28.8	4.1

Using the rules in Scenario 5.4.3, this approach agreed on a vacant status 53.8 percent and 54.9 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about the same as the true positive rate seen in the production scenario. This result show that this approach would have agreed on an unoccupied status 69.1 percent and 74.0 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. Finally, this approach would have assigned an AR vacant status to an NRFU occupied unit 30.9 percent and 26.0 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively.

Table 62: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	398	30.9	4.4	53.8	4.1	15.3	4.8
Reduced Contacts without AR	530	26.0	3.6	54.9	7.0	19.1	7.6

This part documents how often our AR assignment of an occupied status matched to the NRFU field status, as well as how often it did not match. Notice that there were 4,230 and 4,516 cases respectively analyzed in this table. This was less than the number of occupied cases that were removed in production, because of the additional restriction that the 2013 IRS 1040 household must match the 2012 IRS 1040 household. Table 63 shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 63: Resolved Housing Unit Status of Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	4,230	89.5	0.8	10.5	0.8
Reduced Contacts without AR	4,516	83.6	1.5	16.4	1.5

Using the rules in Scenario 5.4.3, the results show agreement on occupancy status 94.2 percent and 94.9 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about the same to the true positive rate seen in the production scenario. This processing assigned an AR occupied status to a NRFU unoccupied unit 5.8 percent and 5.1 percent of the time in the Control Panel and the Reduced Contacts without AR Panel. These differences can be thought of as false positives, i.e., the cases for which were assigned an occupied status based on administrative records, but were unoccupied.

For AR occupied cases, the true positive rates increased slightly; the false positive rates decreased slightly. One conclusion is that possibly we can rely on the most recent year of IRS data instead of requiring two years of matched data. The upshot to this is the amount of identification gained by only relying upon one year of data to remove cases.

Table 64: Housing Unit Status for Resolved Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	3,785	94.2	0.6	5.0	0.5	0.8	0.4
Reduced Contacts without AR	3,776	94.9	0.6	4.6	0.6	0.5	0.2

This alternate identification scenario would have had the following implications across the Control Panel and the Reduced Contacts without AR combined:

- 34.0 percent reduction in the total number of cases identified based on AR
- 11.4 percent increase in the number of cases flagged as AR vacant
- 37.7 percent reduction in the number of cases flagged as AR occupied

- Under the full AR removal approach, the overall reduction in flagged cases would have increased the NRFU workload by 48.4 percent.
- A 54.4 percent true positive vacant rate. In production, it was 53.8 percent
- A 94.6 percent true positive occupied rate. In production, it was 94.1 percent.

The following two tables show the levels used to calculate the percent change metrics just mentioned. They also show the results for the Control Panel and the Reduced Contacts without AR separately.

Table 65: Percent Change in AR Removals by Panel

Panel	AR Occupied		% Change	AR Vacant		% Change	Total AR Removals		% Change
	Production	Alt Scenario		Production	Alt Scenario		Production	Alt Scenario	
Control	7,028	4,230	39.8%	462	512	10.8%	7,490	4,742	36.7%
Reduced Contacts without AR	7,002	4,516	35.5%	665	744	11.9%	7,667	5,260	31.4%
Total	14,030	8,746	37.7%	1,127	1,256	11.4%	15,157	10,002	34.0%

Table 66: Percent Change in NRFU Workload by Panel

Panel	NRFU Universe	Total AR Removals		NRFU Workload*		% Change
		Production	Alt Scenario	Production	Alt Scenario	
Control	13,253	7,490	4,742	5,763	8,511	47.7%
Reduced Contacts without AR	12,553	7,667	5,260	4,886	7,293	49.3%
Total	25,806	15,157	10,002	10,649	15,804	48.4%

*NRFU Workload = NRFU Universe – Total AR Removals. This assumes a full AR removal approach.

5.4.4 What if additional AR sources to remove cases were used?

The 2014 Census Test used the presence of the 2013 IRS 1040, the 2012 and 2013 Medicare data, and the absence of certain UAA reason codes to assign cases as occupied. A question was whether adding additional AR sources would further reduce the workload. Recall that 59.8 percent of the NRFU universe was identified in production.

For this what-if scenario, the Selective Service System Registration, Department of Housing and Urban Development Tenant Rental Assistance Certification System, and Indian Health Service Patient database files were included as additional sources for identifying AR occupied cases. The rules for forming AR households from these sources were applied in a similar manner to those used for the IRS source. That is, if the unit had any PIKs from any of the three sources, no non-PIKed persons were associated with the unit, and none of the applicable UAA reason codes applied, then the unit was flagged as AR occupied. In doing so, 60.8 percent would be identified. Table 67 provides the distribution of initial flags for the Control Panel and the Reduced Contacts without AR Panel.

Table 67: AR Flags by Panel: Alternate AR Removal Scenario Using the Selective Service, HUD TRACS and Indian Health Service Files

Panel	# Units N	Initial Flags					
		No AR Identification		AR Occupied		AR Vacant	
		%	SE	%	SE	%	SE
Control	13,253	42.6	3.4	54.0	3.2	3.4	0.6
Reduced Contacts without AR	12,553	37.8	1.6	57.1	1.7	5.1	0.7

This part documents how often our AR assignment of a vacant status matched to the NRFU field status, as well as how often it did not match. Notice that there were 450 and 646 cases respectively analyzed in this table. This was fewer than the number of vacant cases that were removed in production, because additional AR sources to flag units as occupied were added. Table 68 shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 68: Resolved Housing Unit Status for Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	450	79.6	2.4	20.4	2.4
Reduced Contacts without AR	646	71.2	4.4	28.8	4.4

Using the rules in Scenario 5.4.4, this approach would have agreed on a vacant status 53.1 percent and 53.7 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about the same as the true positive rate seen in the production scenario. This approach would have agreed on an unoccupied status 70.1 percent and 75.7 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This approach would have assigned an AR vacant status to an NRFU occupied unit 29.9 percent and 24.3 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively.

Table 69: Housing Unit Status for Resolved Cases Assigned an AR Vacant Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	358	29.9	4.3	53.1	4.4	17.0	5.2
Reduced Contacts without AR	460	24.3	3.7	53.7	8.1	22.0	8.6

This part documents how often our AR assignment of an occupied status matched to the NRFU field status, as well as how often it did not match. Notice that there are 7,160 and 7,163 cases respectively analyzed in this table. This was near the number of occupied cases that were removed in production. Table 70 below shows the resolved housing unit status distribution for the Control Panel and the Reduced Contacts without AR Panel.

Table 70: Resolved Housing Unit Status of Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Resolved		Unresolved	
		%	SE	%	SE
Control	7,160	89.1	0.9	10.9	0.9
Reduced Contacts without AR	7,163	82.0	1.5	18.0	1.5

Using the rules in Scenario 5.4.4, this approach would have agreed on an occupied status 93.6 percent and 94.3 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. This was about the same as the true positive rate seen in the production scenario. This approach assigned an AR occupied status to a NRFU unoccupied unit 6.4 percent and 5.7 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. For AR vacant and AR occupied cases, the true positive and false positive rates essentially remained the same. This was because the workload only decreased by 1 percent. The addition of these three additional sources does not really change the results in comparison to just using the IRS and Medicare as sources.

Table 71: Housing Unit Status of Resolved Cases Assigned an AR Occupied Flag by Panel

Panel	# Units N	Occupied		Vacant		Non-Existent	
		%	SE	%	SE	%	SE
Control	6,380	93.6	0.6	5.7	0.6	0.7	0.3
Reduced Contacts without AR	5,876	94.3	0.6	5.1	0.6	0.6	0.2

This alternate identification scenario would have had the following implications across the Control Panel and the Reduced Contacts without AR combined:

- 1.7 percent increase in the total number of cases identified based on AR
- 2.8 percent reduction in the number of cases flagged as AR vacant
- 2.1 percent increase in the number of cases flagged as AR occupied
- Under the full AR removal approach, the overall increase in flagged cases would have reduced the NRFU workload by 2.5 percent.
- A 53.4 percent true positive vacant rate. In production, it was 53.8 percent
- A 93.9 percent true positive occupied rate. In production, it was 94.1 percent.

Tables 72 and 73 show the levels used to calculate the percent change metrics just mentioned. They also show the results for the Control Panel and the Reduced Contacts without AR separately.

Table 72: Percent Change in AR Removals by Panel

Panel	AR Occupied		% Change	AR Vacant		% Change	Total AR Removals		% Change
	Production	Alt Scenario		Production	Alt Scenario		Production	Alt Scenario	
Control	7,028	7,160	1.9%	462	450	2.6%	7,490	7,610	1.6%
Reduced Contacts without AR	7,002	7,163	2.3%	665	646	2.9%	7,667	7,809	1.9%
Total	14,030	14,323	2.1%	1,127	1,096	2.8%	15,157	15,419	1.7%

Table 73: Percent Change in NRFU Workload by Panel

Panel	NRFU Universe	Total AR Removals		NRFU Workload*		% Change
		Production	Alt Scenario	Production	Alt Scenario	
Control	13,253	7,490	7,610	5,763	5,643	2.1%
Reduced Contacts without AR	12,553	7,667	7,809	4,886	4,744	2.9%
Total	25,806	15,157	15,419	10,649	10,387	2.5%

*NRFU Workload = NRFU Universe – Total AR Removals. This assumes a full AR removal approach.

6. Related Evaluations, Experiments, and/or Assessments

Related evaluations, experiments and assessments include the 2013 Census Test Assessment, the 2010 Census Match Study, the June 2013 Program Management Review Presentation on Administrative Record Usage in NRFU and the 2014 Census Test Panel Comparison analysis.

7. Dependencies

The Administrative Records Modeling project is dependent on the iterative testing process for the research and testing phase. The results from the 2013 Census Test fed the requirements for the 2014 Census Test. The results from the 2014 Census Test resulted in helping determine the design for the 2015 testing activities for NRFU research. The result of the upcoming 2015 Census Test will be one of the factors in determining the AR usage during NRFU for the 2020 design decision.

8. Conclusions and Recommendations

8.1 Conclusions

For cases that were determined to be AR vacant, the results show agreement on a vacant status 53.3 percent and 54.2 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. Unoccupied status agreed 69.9 percent and 75.4 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. The processing assigned an AR vacant status to an NRFU occupied unit 30.1 percent and 24.6 percent of the time in the Control Panel and the Reduced Contacts without AR Panel, respectively. When looking at these occupancy status case disagreements, the results show 19 percent of control panel cases had either conflicting answers in the interviews or unresolved population counts.

For AR occupied cases, the results show agreement of occupancy status over 90 percent across the different panels. While this was high, the count agreement for the Control Panel was only 54 percent. Further analysis showed that the count agreement was higher for AR single adult with no children and two adults with and without children household compositions. One finding is to give priority to these household compositions when determining when to use ARs.

Our last part of analysis looked at several “what if” scenarios. Using both the Control Panel and the Reduced Contacts without AR Panel, we saw the following changes for different scenarios.

- The first “what-if” scenario showed a 58.4 percent true positive rate when identifying AR vacant units if the identification used the only the vacant reason code to identify AR vacant cases. In production, this value was only 53.8 percent when the vacant and additional codes were used. Under a full AR removal approach, the NRFU workload would have been reduced by 1.8 percent.
- The second “what-if” scenario showed a 93.2 percent true positive rate when using only three AR household compositions to identify AR occupied cases. In production, this value was 94.1 percent. The results show a 64.6 percent population match rate between true positive occupied cases in this scenario, compared to just a 56.1 percent agreement in production. Lastly, under a full AR removal approach, the reduction in identified AR units would amount to an increase of 36.4 percent in the NRFU workload.
- The final “what if” scenario showed a 93.9 percent true positive rate when using additional sources to identify AR occupied cases. In production, this value was 94.1 percent. The results show that under a full AR removal strategy, a reduction in the NRFU workload by 2.5 percent.

8.2 Program-Level Recommendations

The NRFU operation is the single most expensive component of the Census. Research has shown that using modeling and AR data to remove cases from the NRFU workload promises substantial workload reduction for the 2020 Census. The 2020 Census design should include a strategy involving the use of modeling and AR data to remove vacant and delete cases from the initial NRFU workload, and then occupied cases either from the initial workload or after a single contact attempt. The Census Bureau should continue to explore identifying additional sources to better help identify vacant and occupied units.

8.3 Project-Level Recommendations

The 2014 Census Test was implemented using additional UAA Nixie reasons, above and beyond vacant, to identify vacant housing units. This report was able to quantify the additional workload removal and impact on census quality for this based on the rule-based approaches tested here. The recommendation is to continue to research these additional reasons to see if improvements can be made.

The 2014 Census Test implemented a rule-based approach to identify AR occupied units. This research has shown the AR household compositions based on a) one adult with no children present, b) two adult with no children present and c) two adult with children present had higher match rates to the census counts than the remaining categories. One recommendation is to see if these household composition categories can be utilized more fully in future research. Another recommendation is to continue to research predictive modeling approaches in addition to rule-based used in this analysis.

9. Knowledge Management Resolutions

The Administrative Records Modeling team has several Knowledge Management recommendations related to using AR information during the NRFU operation to reduce or determine when to stop the data collection workload. The 2014 Census Test allowed us to gather more information about the usage and potential census quality implications. This material, along with analysis using 2010 Census data and the upcoming 2015 Census Test, will allow us to be able to fully evaluate the knowledge management items that we have been asked to investigate.

10. Acknowledgements

Thanks to the Administrative Records Modeling team, the 2014 Census Test Team, Deborah Wagner and her staff in the Center for Administrative Records Research and Applications.

11. References

Bishop, Deirdre D. (2014) “The Path to the 2020 Census Design Decision.” U.S. Census Bureau

Keller, A. and T. Fox (2012) “2010 Census Coverage Measurement Estimation Report: Components of Census Coverage for the Household Population in the United States.” DSSD 2010 CENSUS COVERAGE MEASUREMENT MEMORANDUM SERIES #2010-G-04. U.S. Census Bureau. Washington, DC.

Poehler et al (2014) “2020 Research and Testing: 2014 Census Test Nonresponse Followup Panel Comparisons and Instrument Analysis.” U.S. Census Bureau

Rastogi and O'Hara (2012) – Census 2010 Match Study U.S. Census Bureau

Walejko, Gina K., Andrew Keller, Gianna Dusch, and Peter V. Miller (2014) . “2020 Research and Testing: 2013 Census Test Assessment.” U.S. Census Bureau

12. Document Logs

12.1 Sensitivity Assessment

This table specifies whether the document contains any administratively restricted information.

Verification of Document Content
<p>This document does not contain any:</p> <ul style="list-style-type: none"> • Title 5, Title 13, Title 26, or Title 42 protected information; • Procurement information; • Budgetary information; and/or, • Personally identifiable information.

Verification of Document Content

Document Authors: Andrew Keller, Tyler Fox and Thomas Mule	Date: 7/8/2015
---------------------------------------------------------------	-------------------

12.2 Review/Approval

This table documents the review level and/or approval authority.

Document Review and Approval Tier: Program Document		
Name	Area Represented	Date

12.3 Version History

The document version history recorded in this section provides the revision number, the version number, the date it was issued, and a brief description of the changes since the previous release. Baseline releases are also noted.

Version	Date	Description
V1.0	12/10/14	Draft for Decennial Leadership Group presentation and Program Management Review.
V1.1	4/3/15	Draft for baseline submission
V1.2	6/04/15	Draft based on baseline submission comments
V1.3	7/8/15	Draft updated based on final comments
V1.4	10/29/15	Updated for signature review comments
V1.5	1/19/16	Updated for final signature review comments